



(12) 发明专利

(10) 授权公告号 CN 108492000 B
(45) 授权公告日 2021.07.02

(21) 申请号 201810141410.6
(22) 申请日 2018.02.11
(65) 同一申请的已公布的文献号
申请公布号 CN 108492000 A
(43) 申请公布日 2018.09.04
(73) 专利权人 浙江大学
地址 310058 浙江省杭州市西湖区余杭塘路866号
(72) 发明人 赵春晖 胡赞昀
(74) 专利代理机构 杭州求是专利事务所有限公司 33200
代理人 黄欢娣 邱启旺
(51) Int.Cl.
G06Q 10/06 (2012.01) (续)
(56) 对比文件
CN 105911977 A, 2016.08.31
CN 106897542 A, 2017.06.27
US 2006074598 A1, 2006.04.06
US 2011218737 A1, 2011.09.08
CN 105204493 A, 2015.12.30
CN 104699077 A, 2015.06.10

CN 106680012 A, 2017.05.17
CN 107633249 A, 2018.01.26
CN 106895975 A, 2017.06.27
US 2013090068 A1, 2013.04.11
陈昭 等. 时序非平稳性ADF检验法的理论与应用.《广州大学学报(自然科学版)》.2008, 第7卷(第5期), 第5-10页.
石海忱 等. 协整系数矩阵的非平稳工程系统故障诊断应用研究.《振动与冲击》.2015, 第34卷(第1期), 第145-150页.
Q. Peter He et al.. A New Fault Diagnosis Method Using Fault Directions in Fisher Discriminant Analysis.《American Institute of Chemical Engineers》.2005, 第555-571页.
Yue Wang et al.. Probabilistic fault diagnosis method based on the combination of nest-loop fisher discriminant analysis and analysis of relative changes.《Control Engineering Practice》.2017, 第32-45页. (续)

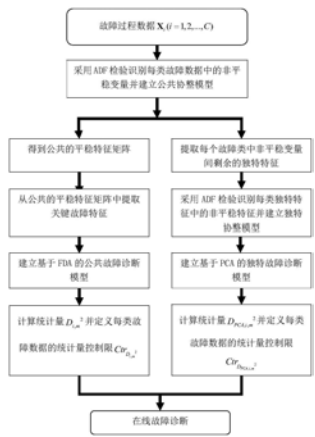
审查员 任丽娜

权利要求书5页 说明书12页 附图6页

(54) 发明名称
面向百万千瓦超超临界机组非平稳特性的故障诊断方法

临界发电机组的安全可靠运行并提高了生产效率。

(57) 摘要
本发明公开了一种面向百万千瓦超超临界机组非平稳故障特性的在线故障诊断方法。本发明针对百万千瓦超临界机组故障过程非平稳、复杂多变的特性, 结合运用协整分析和基于变量选择的双重模型用于故障诊断。该方法克服了大型燃煤机组故障过程的非平稳特性, 充分挖掘故障间所包含的潜在信息, 建立了双重诊断模型, 互为补充, 相辅相成, 有效地解决了复杂非平稳故障过程在线诊断困难的问题, 大大提高了非平稳故障过程在线诊断的性能, 有助于现场工程师对故障进行准确的修复, 从而保证了百万千瓦超超



CN 108492000 B

[接上页]

(51) Int.Cl.

G06Q 50/06 (2012.01)

(56) 对比文件

He Sun et al..A Sparse Reconstruction

Strategy for Online Fault Diagnosis in Nonstationary Processes with No a Priori Fault Information.《Industrial & Engineering Chemistry Research》.2017,第6993-7008页.

1.一种面向百万千瓦超超临界机组非平稳故障特性的在线故障诊断方法,其特征在于,该方法包括以下步骤:

(1) 获取过程分析数据:设一个热力系统生产过程具有J个测量变量和操作变量,则每一次采样得到一个 $1 \times J$ 的向量,采样N次后得到的数据表述为一个二维矩阵 $X(N \times J)$,所述测量变量为运行过程中可被测量的状态参数,包括流量、温度、速率;所述操作变量包括进风量、给料量、阀门开度;分别获取C类故障数据二维矩阵,统一表示为 $X_i(N \times J)$,其中下标i表示故障的类别;

(2) 建立公共的故障诊断模型,该步骤通过以下子步骤来实现:

(2.1) 平稳性检验:分别对每个故障类的所有变量进行ADF检验,得到的平稳变量矩阵表示为 $X_{s,i}(N \times J_{s,i})$,非平稳变量矩阵表示为 $X_{ns,i}(N \times J_{ns,i})$,其中,下标s表示平稳变量,ns表示非平稳变量,i表示故障的类别,括号里的 $J_{s,i}$ 表示平稳变量的个数, $J_{ns,i}$ 表示非平稳变量个数;

(2.2) 建立公共的协整模型,该步骤由以下子步骤来实现:

(2.2.1) 获取公共平稳变量矩阵:通过(2.1)区分了每个故障类中的平稳和非平稳变量,求取所有故障类中平稳变量的交集,得到对于每个故障类都平稳的变量,组成公共平稳变量矩阵:

$$X_s(N \times J_s) = X_{s,1} \cap X_{s,2} \cap \dots \cap X_{s,C} \quad (1)$$

其中, $\mathbf{X}_{s,i} = [\mathbf{x}_{s,1}, \mathbf{x}_{s,2}, \dots, \mathbf{x}_{s,j}, \dots, \mathbf{x}_{s,J_{s,i}}]$ $j \in (1, 2, \dots, J_{s,i})$, $\mathbf{x}_{s,j}$ 代表第i个故障类中的平稳变量, $\cap$ 表示求取不同矩阵的交集, J_s 表示综合C类故障后的公共平稳变量个数;

(2.2.2) 获取所有非平稳变量:求取所有故障类中非平稳变量的并集,得到C种故障类型下的全部非平稳变量集合:

$$X_{ns}(N \times J_{ns}) = X_{ns,1} \cup X_{ns,2} \cup \dots \cup X_{ns,C} \quad (2)$$

其中, $\mathbf{X}_{ns,i} = [\mathbf{x}_{ns,1}, \mathbf{x}_{ns,2}, \dots, \mathbf{x}_{ns,j}, \dots, \mathbf{x}_{ns,J_{ns,i}}]$ $j \in (1, 2, \dots, J_{ns,i})$, $\mathbf{x}_{ns,j}$ 表示第i个故障类中的非平稳变量, $\cup$ 表示求取不同矩阵的并集, J_{ns} 表示综合C类故障后的所有非平稳变量个数;

(2.2.3) 建立公共协整模型:对(3.2)中的所有非平稳变量矩阵 $X_{ns}(N \times J_{ns})$ 进行协整分析,得到公共协整向量B以及如式(3)所示的公共协整特征矩阵:

$$\gamma = X_{ns} B \quad (3)$$

其中,B是 $J_{ns} \times R_{CA}$ 的矩阵,向量个数 R_{CA} 由Johansen检验确定; $\gamma = X_{ns} B$ 是从所有类的非平稳向量中提取的公共平稳残差序列,表征了公共的故障特性;

(2.3) 获取公共平稳特征矩阵:求取公共平稳变量矩阵和公共协整特征矩阵的并集,得到公共的平稳特征矩阵:

$$X_{alls}(N \times J_{alls}) = X_s \cup \gamma \quad (4)$$

其中, X_{alls} 包含了所有(2.2.1)中的公共平稳变量和(2.2.3)中的公共平稳协整特征,统称为公共平稳特征, J_{alls} 表示公共平稳特征的个数;

(2.4) 提取关键故障特征,该步骤由以下子步骤来实现:

(2.4.1) 假设所有的故障样本服从多元广义高斯分布;定义一个虚拟的故障变量A,取值的个数与故障数据二维矩阵的组数相同,即为C;服从如下的概率分布 $P(A=c) = P(c)$;根据互信息准则, x_{alls} 和故障变量A间的互信息按下式表示:

$$\begin{aligned}
I(\mathbf{x}_{all}; A) &= \sum_{c=1}^C \int_x P(c, x) \log \left(\frac{P(c, x)}{P(c)P(x)} \right) dx \\
&= \sum_{c=1}^C \int_x P(c) P(x|c) \log \left(\frac{P(c)P(x|c)}{P(c)P(x)} \right) dx \\
&= \sum_{c=1}^C P(c) \int_x P(x|c) \log(P(x|c)) dx - \sum_{c=1}^C P(c) \int_x P(x|c) \log(P(x)) dx
\end{aligned} \quad (5)$$

其中, $\mathbf{x}_{all}$ 是公共平稳特征矩阵中的一个随机向量, x 表示其中的一个样本, 样本服从散布矩阵为 Σ , 形状参数为 β 的多元广义高斯分布; $\mathbf{x}_{all}$ 在 $A=c$ 条件下的概率密度服从参数为 Σ_c 和 β_c ;

(2.4.2) 将所有特征按互信息大小递减排序后, 需要找出在所有故障类中分类性能最好的一组关键特征, 表示为 X_{key} ; 该关键特征 $X_{key} = [X_{key,1}, X_{key,2}, \dots, X_{key,i}, \dots, X_{key,C}]$, $X_{key,i}$ 中具有 N_i 个样本;

(2.5) 建立基于FDA的公共故障诊断模型, 该步骤包括以下子步骤:

(2.5.1) 选取上述得到的所有故障类的关键特征中的样本作为总样本 $\mathbf{X}_{key} \left(\sum_{i=1}^C N_i \times J_{key} \right)$, 其中, $\mathbf{X}_{key} \left(\sum_{i=1}^C N_i \times J_{key} \right)$ 由 $X_{key,i}$ ($i=1, 2, \dots, C$) 组成, J_{key} 代表关键故障特征的个数;

(2.5.2) 数据准备: 分别计算总样本均值向量 $\bar{\mathbf{x}}$ 、每类故障样本均值向量 $\bar{\mathbf{x}}_i$, 总类内散布矩阵 S_w 和类间的散布矩阵 S_b , 计算公式如下:

$$\begin{aligned}
S_i &= \sum_{\mathbf{x}_i \in X_{key,i}} (\mathbf{x}_i - \bar{\mathbf{x}}_i)(\mathbf{x}_i - \bar{\mathbf{x}}_i)^T \\
S_w &= \sum_{i=1}^C S_i \\
S_b &= \sum_{\mathbf{x}_i \in X_{key,i}} N_i (\bar{\mathbf{x}}_i - \bar{\mathbf{x}})(\bar{\mathbf{x}}_i - \bar{\mathbf{x}})^T
\end{aligned} \quad (6)$$

其中, S_i 是每个故障类的散布矩阵;

(2.5.3) 最大化费舍尔准则: 求取使费舍尔准则函数最大的权重向量 $\mathbf{w}$, 即求使类间离散度和类内离散度的比值最大的权重向量 $\mathbf{w}$:

$$J(\mathbf{w}) = \frac{\mathbf{w}^T S_b \mathbf{w}}{\mathbf{w}^T S_w \mathbf{w}} \quad (7)$$

其中, 权重向量 $\mathbf{w}$ 就是费舍尔最优判别方向;

(2.5.4) 求使费舍尔判据最大的权重向量相当于一个求取广义特征向量的特征值问题:

$$S_b \mathbf{w} = \lambda_1 S_w \mathbf{w} \quad (8)$$

其中, 特征值 λ_1 代表故障类之间的总体可分程度; 当 S_w 为非奇异阵时, 上式转化成一个常规的特征值问题:

$$S_w^{-1} S_b \mathbf{w} = \lambda_1 \mathbf{w} \quad (9)$$

求解上式得到 R_{FDA} 个非零的广义特征值以及广义特征向量矩阵 $R^*(J_{key} \times R_{FDA})$, $R_{FDA} = C -$

1;

(2.5.5) 计算每类故障样本的类内判别成分和故障衡量统计指标 $D_{i,m}^2$:

$$T_i = X_{key,i} R^*$$

$$D_{i,m}^2 = (T_i - \bar{T}_i)^T \Sigma_i^{-1} (T_i - \bar{T}_i) \quad (10)$$

其中, T_i 表示判别成分矩阵, $\bar{T}_i$ 是 T_i 的均值向量, Σ_i 是 T_i 的协方差矩阵;

(2.5.6) 建立每类故障特征衡量指标的控制限;故障衡量统计指标 $D_{i,m}^2$ 服从加权 χ^2 分布,则依据带权重的 χ^2 分布建立统计量的控制限 $Ctr_{D_{i,m}^2}$;

(3) 建立独特的故障诊断模型,该步骤通过以下子步骤来实现:

(3.1) 利用式(11)提取剩余的协整关系:

$$B^\perp = [I - B(B^T B)^{-1} B^T] \quad (11)$$

其中, B 是从(2.2.3)中提取的公共协整向量矩阵, I 表示单位矩阵; $B^\perp$ 是 B 的正交空间,表示剩余的协整方向;

(3.2) 将每个故障类中非平稳变量 $\tilde{X}_{ns,i} (N \times J_{ns})$ 投影到剩余的独特子空间中,得到剩余的独特特征:

$$X_{\perp,i} = \tilde{X}_{ns,i} B^\perp \quad (12)$$

其中, $\tilde{X}_{ns,i}$ 表示 X_{ns} 中的第 i 类样本矩阵;

(3.3) 平稳性检验:

分别对(3.2)得到的每个故障类中所有非平稳变量间剩余的独特特征为剩余独特特征,进行ADF检验,得到的平稳的剩余独特特征矩阵表示为 $X_{s\perp,i}$,非平稳的剩余独特特征矩阵表示为 $X_{ns\perp,i}$;

(3.4) 建立每个故障类独特的协整模型:

对(3.3)中的非平稳的剩余独特特征矩阵 $X_{ns\perp,i}$ 进行协整分析,得到独特的协整向量 $B_{specific,i}$ 以及如式(13)所示的独特的协整特征:

$$\gamma_{specific,i} = X_{ns\perp,i} B_{specific,i} \quad (13)$$

其中, $B_{specific,i}$ 表示第 i 个故障类的独特的协整向量,向量个数由Johansen检验确定;

(3.5) 获取独特的平稳特征矩阵:求取每个故障类平稳的剩余独特特征矩阵和独特的协整特征矩阵的并集,得到每个类独特的平稳特征矩阵:

$$X_{specific,i} = X_{s\perp,i} \cup \gamma_{specific,i} \quad (14)$$

其中, $X_{specific,i}$ 包含了所有(3.3)中的平稳的剩余独特特征和(3.4)中的独特的协整特征,统称为独特的平稳特征;

(3.6) 为每个故障类分别建立基于PCA的独特故障诊断模型,该步骤包括以下子步骤:

(3.6.1) 选取上述得到的每个故障类的独特的平稳特征的样本作为总样本;

(3.6.2) 求取PCA的负载向量相当于求取总样本协方差矩阵的特征向量:

$$S_{\Sigma,i} = X_{specific,i}^T X_{specific,i} \quad (15)$$

$$S_{\Sigma,i} P_i = \lambda_2 P_i$$

其中, $S_{\Sigma,i}$ 表示总样本的协方差矩阵;这样得到的负载矩阵 P_i 就是PCA的最佳投影方向;

(3.6.3) 计算每类故障样本的主成分和故障衡量统计指标 $D_{PCA,i,m}^2$:

$$T_{PCA,i} = X_{specific,i} P_i$$

$$D_{PCA,i,m}^2 = (T_{PCA,i} - \bar{t}_{PCA,i})^T \Sigma_{PCA,i}^{-1} (T_{PCA,i} - \bar{t}_{PCA,i}) \quad (16)$$

其中, $T_{PCA,i}$ 表示主成分矩阵,包含了大部分的样本波动信息, $\bar{t}_{PCA,i}$ 是 $T_{PCA,i}$ 的均值向量, $\Sigma_{PCA,i}$ 是 $T_{PCA,i}$ 的协方差矩阵;

(3.6.4) 建立每类故障特征衡量指标的控制限;由于假设过程数据服从多元广义高斯分布,且样本数量足够大,故认为变量衡量指标 $D_{PCA,i,m}^2$ 服从加权 χ^2 分布,则依据带权重的 χ^2 分布建立统计量的控制限 $Ctr_{D_{PCA,i,m}^2}$;

(4) 在线故障诊断:当故障发生后,需要检测发生故障的变量,确定故障的类型,该步骤包括以下子步骤:

(4.1) 按照步骤(1)获取新数据 $\tilde{x}_f(1 \times J)$,根据步骤(2.2)到(2.3),得到新的公共平稳特征矩阵 $\tilde{x}_{new}(1 \times J_{alls})$;

(4.2) 按照步骤(2.4)提取出关键故障特征,得到新数据的关键故障特征样本 $\tilde{x}_{key}(1 \times J_{key})$;

(4.3) 调用公共故障诊断模型:

$$\tilde{t}_{new} = \tilde{x}_{key} R^*$$

$$\tilde{D}_{new,i}^2 = (\tilde{t}_{new} - \bar{t}_i)^T \Sigma_i^{-1} (\tilde{t}_{new} - \bar{t}_i) \quad (17)$$

其中, $R^*(J_{key} \times R_{FDA})$ 是步骤(2.5.4)中计算得到的FDA的广义特征向量矩阵, $\tilde{D}_{new,i}^2$ 为公共故障衡量指标, $\tilde{x}_{key}$ 是新数据的关键故障特征, $\tilde{t}_{new}$ 是 $\tilde{x}_{key}$ 的判别成分, $\bar{t}_i$ 和 Σ_i 均由步骤(2.5.5)求得;

(4.4) 按步骤(3.2)将新的样本投影到剩余的独特子空间中,得到 $\tilde{x}_{\perp,new}$;将 $\tilde{x}_{\perp,new}$ 代入到每个故障类的独特协整模型中,得到新数据的独特平稳特征样本 $\tilde{x}_{\perp,new,i}$;

(4.5) 调用独特故障诊断模型:

$$\tilde{t}_{\perp,new,i} = \tilde{x}_{\perp,new,i} P_i$$

$$\tilde{D}_{\perp,new,i}^2 = (\tilde{t}_{\perp,new,i} - \bar{t}_{PCA,i})^T \Sigma_{PCA,i}^{-1} (\tilde{t}_{\perp,new,i} - \bar{t}_{PCA,i}) \quad (18)$$

其中, P_i 是步骤(3.6.2)中计算得到的PCA的负载矩阵, $\tilde{x}_{\perp,new,i}$ 是新数据的独特平稳特征样本, $\tilde{t}_{\perp,new,i}$ 是 $\tilde{x}_{\perp,new,i}$ 的主成分矩阵, $\bar{t}_{PCA,i}$ 和 $\Sigma_{PCA,i}$ 均由步骤(3.6.3)求得, $\tilde{D}_{\perp,new,i}^2$ 为独特故障指标;

(4.6) 对每类候选故障分别比较 $\tilde{D}_{new,i}^2$ 和其控制限 $Ctr_{D_{i,m}^2}$, $\tilde{D}_{\perp,new,i}^2$ 和其控制限 $Ctr_{D_{PCA,i,m}^2}$,若新来的样本在该类候选故障的故障诊断模型下均没有超限报警,则表明机组发生了该候选故障。

2. 根据权利要求1所述的方法,其特征在于,所述步骤(2.1)区分了每个故障类中的平

稳和非平稳变量。

3. 根据权利要求1所述的方法, 其特征在于, 由于假设过程数据服从多元广义高斯分布, 且样本数量足够大, 故认为变量衡量指标 $D_{i,m}^2$ 服从加权 χ^2 分布。

4. 根据权利要求1所述的方法, 其特征在于, 2.2.1中, $\cap$ 表示求取不同矩阵的交集。

面向百万千瓦超超临界机组非平稳特性的故障诊断方法

技术领域

[0001] 本发明属于非平稳过程统计监测领域,特别是针对一种面向百万千瓦超超临界机组非平稳故障特性的在线故障诊断方法。

背景技术

[0002] 近年来,为了实现电力的可持续发展,火力发电行业积极开展结构调整,以大容量、高参数、低能耗的超临界和超超临界机组取代高能耗小型火电机组,基本形成了以超超临界机组为主体的电力能源结构。百万千瓦超超临界机组热力系统是个极其复杂的工业过程,规模庞大、设备多数、参数众多且相互影响,现场具有高温、高压、高噪声等特点。

[0003] 整个工艺过程生产流程长、单元装置多、空间分布广、安全要求高,其中可能发生的故障复杂多样。以锅炉及其辅助设备为例,常见的典型故障包括由于磨损、腐蚀、设备老化、应力拉伸等原因造成的四管泄露(水冷壁、过热器、再热器、省煤器),由于堵塞、漏粉、断裂等原因造成的磨煤机故障,由于振动大、温度高等原因造成的送风机、引风机故障等。为了保障机组安全可靠运行,必须对热力系统采用有效的故障检测和故障诊断方法。故障诊断是指在检测到故障后,进一步判断发生了哪种故障,并及时对故障变量进行隔离,消除故障的影响。

[0004] 前人对此已经作了相应的研究与探讨,基于不同的角度提出了相应的在线故障诊断方法。概括说来,主要包括基于解析模型的方法和基于数据驱动的方法。对于火电机组这样的复杂非线性系统,其结构复杂、动态时变以及强耦合性严重,使得难以建立精确的解析模型来实现故障诊断。而随着技术的发展,电厂生产过程中可以获得越来越多的数据,基于数据驱动的在线故障诊断策略越来越受到研究人员和现场工程师的青睐。基于数据、驱动的方法主要包括多元统计方法和人工智能方法。其中,主成分分析、偏最小二乘法、费舍尔判别分析等方法作为多元统计分析方法的典型代表,受到了大家的关注。其原理是假设机组在预先设定的稳态工况下运行,一旦运行状态偏离稳态工况,则认为机组发生了故障。并通过两个监测统计量 T^2 和SPE来判断机组是否偏离正常状态。但是,基于传统多元统计分析在线故障诊断方法应用于实际超超临界百万机组的生产过程中存在着两方面的问题:首先,百万千瓦超超临界机组在运行过程中可能发生很多故障情况。受到内部和外部因素的影响,故障过程中的变量和输出会随着时间的推移而变化,从而使整个故障过程具有非平稳特性。由于非平稳过程的统计特性会随时间变化而变化,传统的多元统计分析方法无法准确描述非平稳过程,因而难以用于非平稳过程的建模和诊断。其次,超超临界机组热力系统结构复杂,参数众多且相互影响,故障特性复杂多变。传统的故障诊断方法通常只利用正常数据与故障数据的差别建立监测指标,而没有考虑不同故障间的关系。针对上述问题,前人提出将所有故障考虑在内,利用不同故障类的数据进行分析。但他们的做法局限于仅考虑不同故障类数据在分布和方差信息上的差异,将多个故障的诊断视为了一个简单的多分类问题,而没有考虑不同故障类间更深层的差异和联系。因而,现有的故障诊断方法并不能很好地运用在百万千瓦超超临界机组上。

[0005] 本发明的内容深入考虑了百万千瓦超超临界机组非平稳以及复杂多样的故障特性,提出了一种基于协整分析和双重模型的故障诊断策略。该方法立足于故障过程的非平稳特性,区分了平稳变量和非平稳变量。并充分挖掘了不同的故障类间的公共信息和每个故障类的独特信息,分别建立了公共故障诊断模型和独特故障诊断模型,综合考虑这两个模型进行判断,大大提高了复杂非平稳过程在线故障诊断的性能。到目前为止,尚未见到与本发明相关的研究报道。

发明内容

[0006] 本发明的目的在于针对百万千瓦超超临界机组的复杂非平稳故障特性,提供了一种面向百万千瓦超超临界机组非平稳故障特性的在线故障诊断方法。

[0007] 本发明的目的是通过以下技术方案来实现的:提供一种面向百万千瓦超超临界机组的非平稳故障特性分析与在线故障诊断方法,该方法包括以下步骤:

[0008] (1) 获取过程分析数据:设一个热力系统生产过程具有J个测量变量和操作变量,则每一次采样可以得到一个 $1 \times J$ 的向量,采样N次后得到的数据表述为一个二维矩阵 $X(N \times J)$,所述测量变量为运行过程中可被测量的状态参数,包括流量、温度、速率;所述操作变量包括进风量、给料量、阀门开度;分别获取C组故障数据二维矩阵,统一表示为 $X_i(N \times J)$,其中下标i表示故障的类别;

[0009] (2) 建立公共的故障诊断模型,该步骤通过以下子步骤来实现:

[0010] (2.1) 平稳性检验:分别对每个故障类的所有变量进行ADF检验,得到的平稳变量矩阵表示为 $X_{s,i}(N \times J_{s,i})$,非平稳变量矩阵表示为 $X_{ns,i}(N \times J_{ns,i})$,其中,下标s表示平稳变量,ns表示非平稳变量,i表示故障的类别,括号里的 $J_{s,i}$ 表示平稳变量的个数, $J_{ns,i}$ 表示非平稳变量个数。

[0011] (2.2) 建立公共的协整模型,该步骤由以下子步骤来实现:

[0012] (2.2.1) 获取公共平稳变量矩阵:通过(2.1)区分了每个故障类中的平稳和非平稳变量,求取所有故障类中平稳变量的交集,可以得到对于每个故障类都平稳的变量,组成公共平稳变量矩阵:

$$[0013] \quad X_s(N \times J_s) = X_{s,1} \cap X_{s,2} \cap \dots \cap X_{s,C} \quad (1)$$

[0014] 其中, $X_{s,i} = [x_{s,i,1}, x_{s,i,2}, \dots, x_{s,i,j}, \dots, x_{s,i,J_{s,i}}]$ $j \in (1, 2, \dots, J_{s,i})$, $x_{s,i,j}$ 代表第i个故障类中的平稳变量, $\cap$ 表示求取不同矩阵的交集, J_s 表示综合C类故障后的公共平稳变量个数。

[0015] (2.2.2) 获取所有非平稳变量:求取所有故障类中非平稳变量的并集,可以得到C种故障类型下的全部非平稳变量集合:

$$[0016] \quad X_{ns}(N \times J_{ns}) = X_{ns,1} \cup X_{ns,2} \cup \dots \cup X_{ns,C} \quad (2)$$

[0017] 其中, $X_{ns,i} = [x_{ns,i,1}, x_{ns,i,2}, \dots, x_{ns,i,j}, \dots, x_{ns,i,J_{ns,i}}]$ $j \in (1, 2, \dots, J_{ns,i})$, $x_{ns,i,j}$ 表示第i个故障类中的非平稳变量, $\cup$ 表示求取不同矩阵的并集, J_{ns} 表示综合C类故障后的所有非平稳变量个数。

[0018] (2.2.3) 建立公共协整模型:对(3.2)中的所有非平稳变量矩阵 $X_{ns}(N \times J_{ns})$ 进行协整分析,得到公共协整向量B以及如式(3)所示的公共协整特征:

$$[0019] \quad \gamma = X_{ns} B \quad (3)$$

[0020] 其中,B是 $J_{ns} \times R_{CA}$ 的矩阵,向量个数 R_{CA} 由Johansen检验确定。 $\gamma = X_{ns} B$ 是从所有类

的非平稳向量中提取的公共平稳残差序列,表征了公共的故障特性。

[0021] (2.3) 获取公共平稳特征矩阵:求取公共平稳变量矩阵和公共协整特征矩阵的并集,可以得到公共的平稳特征矩阵:

$$[0022] \quad X_{\text{alls}}(N \times J_{\text{alls}}) = X_s \cup \gamma \quad (4)$$

[0023] 其中, X_{alls} 包含了所有 (2.2.1) 中的公共平稳变量和 (2.2.3) 中的公共平稳协整特征,统称为公共平稳特征, J_{alls} 表示公共平稳特征的个数。

[0024] (2.4) 提取关键故障特征,该步骤由以下子步骤来实现:

[0025] (2.4.1) 假设所有的故障样本服从多元广义高斯分布。定义一个虚拟的故障变量 C ,可能的取值的个数与故障数据二维矩阵的组数相同,即为 C 。服从如下的概率分布 $P(C=c) = P(c)$ 。根据互信息准则, x_{alls} 和故障变量 C 间的互信息可以按下式表示

$$[0026] \quad \begin{aligned} I(\mathbf{x}_{\text{all}}; C) &= \sum_{c=1}^C \int_x P(c, x) \log \left(\frac{P(c, x)}{P(c)P(x)} \right) dx \\ &= \sum_{c=1}^C \int_x P(c) P(x|c) \log \left(\frac{P(c)P(x|c)}{P(c)P(x)} \right) dx \\ &= \sum_{c=1}^C P(c) \int_x P(x|c) \log(P(x|c)) dx - \sum_{c=1}^C P(c) \int_x P(x|c) \log(P(x)) dx \end{aligned} \quad (5)$$

[0027] 其中, x_{alls} 是公共平稳特征矩阵中的一个随机向量, x 表示其中的一个样本,样本服从散布矩阵为 Σ , 形状参数为 β 的多元广义高斯分布。 x_{alls} 在 $C=c$ 条件下的概率密度服从参数为 Σ_c 和 β_c 。

[0028] (2.4.2) 将所有特征按互信息大小递减排序后,需要找出那些在所有故障类中分类性能最好的一组关键特征,表示为 X_{key} 。该关键特征 $X_{\text{key}} = [X_{\text{key},1}, X_{\text{key},2}, \dots, X_{\text{key},i}, \dots, X_{\text{key},C}]$, $X_{\text{key},i}$ 中具有 N_i 个样本。

[0029] (2.5) 建立基于FDA的公共故障诊断模型,该步骤包括以下子步骤:

[0030] (2.5.1) 选取上述得到的所有故障类的关键特征中的样本作为总样本

$\mathbf{X}_{\text{key}} \left(\sum_{i=1}^C N_i \times J_{\text{key}} \right)$, 其中, $\mathbf{X}_{\text{key}} \left(\sum_{i=1}^C N_i \times J_{\text{key}} \right)$ 由 $X_{\text{key},i}$ ($i=1, 2, \dots, C$) 组成, J_{key} 代表关键故障特征的个数;

[0031] (2.5.2) 数据准备:分别计算总样本均值向量 $\bar{\mathbf{x}}$ 、每类故障样本均值向量 $\bar{\mathbf{x}}_i$, 总类内散布矩阵 S_w 和类间的散布矩阵 S_b , 计算公式如下:

$$[0032] \quad \begin{aligned} S_i &= \sum_{\mathbf{x}_i \in \mathbf{X}_{\text{key},i}} (\mathbf{x}_i - \bar{\mathbf{x}}_i)(\mathbf{x}_i - \bar{\mathbf{x}}_i)^T \\ S_w &= \sum_{i=1}^C S_i \\ S_b &= \sum_{\mathbf{x}_i \in \mathbf{X}_{\text{key},i}} N_i (\bar{\mathbf{x}}_i - \bar{\mathbf{x}})(\bar{\mathbf{x}}_i - \bar{\mathbf{x}})^T \end{aligned} \quad (6)$$

[0033] 其中, S_i 是每个故障类的散布矩阵;

[0034] (2.5.3) 最大化费舍尔准则:求取使费舍尔准则函数最大的权重向量 w , 即求使类间离散度和类内离散度的比值最大的权重向量 w :

[0035] $J(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{S}_b \mathbf{w}}{\mathbf{w}^T \mathbf{S}_w \mathbf{w}} \quad (7)$

[0036] 其中,权重向量 $\mathbf{w}$ 就是费舍尔最优判别方向。

[0037] (2.5.4) 求使费舍尔判据最大的权重向量相当于一个求取广义特征向量的特征值问题:

[0038] $\mathbf{S}_b \mathbf{w} = \lambda \mathbf{S}_w \mathbf{w} \quad (8)$

[0039] 其中,特征值 λ 代表故障类之间的总体可分程度。当 $\mathbf{S}_w$ 为非奇异阵时,上式可以转化成一个常规的特征值问题:

[0040] $\mathbf{S}_w^{-1} \mathbf{S}_b \mathbf{w} = \lambda \mathbf{w} \quad (9)$

[0041] 求解上式可以得到 R_{FDA} 个非零的广义特征值以及广义特征向量矩阵 $\mathbf{R}^* (\mathbf{J}_{key} \times R_{FDA})$, $R_{FDA} = C - 1$ 。

[0042] (2.5.5) 计算每类故障样本的类内判别成分和故障衡量统计指标 $D_{i,m}^2$:

[0043] $\mathbf{T}_i = \mathbf{X}_{key,i} \mathbf{R}^*$

[0044] $D_{i,m}^2 = (\mathbf{T}_i - \bar{\mathbf{T}}_i)^T \boldsymbol{\Sigma}_i^{-1} (\mathbf{T}_i - \bar{\mathbf{T}}_i) \quad (10)$

[0045] 其中, $\mathbf{T}_i$ 表示判别成分矩阵, $\bar{\mathbf{T}}_i$ 是 $\mathbf{T}_i$ 的均值向量, $\boldsymbol{\Sigma}_i$ 是 $\mathbf{T}_i$ 的协方差矩阵。

[0046] (2.5.6) 建立每类故障特征衡量指标的控制限;由于假设过程数据服从多元广义高斯分布,且样本数量足够大,故认为变量衡量指标 $D_{i,m}^2$ 服从加权 χ^2 分布,则可依据带权重的 χ^2 分布建立统计量的控制限 $Ctr_{D_{i,m}^2}$ 。

[0047] (3) 建立独特的故障诊断模型,该步骤通过以下子步骤来实现:

[0048] (3.1) 利用式(11)提取剩余的协整关系:

[0049] $\mathbf{B}^\perp = [\mathbf{I} - \mathbf{B}(\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T] \quad (11)$

[0050] 其中, $\mathbf{B}$ 是从(2.2.3)中提取的公共协整向量矩阵, $\mathbf{I}$ 表示单位矩阵。 $\mathbf{B}^\perp$ 是 $\mathbf{B}$ 的正交空间,表示剩余的协整方向。

[0051] (3.2) 将每个故障类中非平稳变量 $\tilde{\mathbf{X}}_{ns,i} (N \times J_{ns})$ 投影到剩余的独特子空间中,得到剩余的独特特征:

[0052] $\mathbf{X}_{\perp,i} = \tilde{\mathbf{X}}_{ns,i} \mathbf{B}^\perp \quad (12)$

[0053] 其中, $\tilde{\mathbf{X}}_{ns,i}$ 表示 $\mathbf{X}_{ns}$ 中的第 i 类故障样本矩阵。

[0054] (3.3) 平稳性检验:

[0055] 分别对(3.2)得到的每个故障类中所有非平稳变量间剩余的独特特征(以下简称为每个故障类的剩余独特特征)进行ADF检验,得到的平稳的剩余独特特征矩阵表示为 $\mathbf{X}_{s\perp,i}$,非平稳的剩余独特特征矩阵表示为 $\mathbf{X}_{ns\perp,i}$ 。

[0056] (3.4) 建立每个故障类独特的协整模型:

[0057] 对(3.3)中的非平稳的剩余独特特征矩阵 $\mathbf{X}_{ns\perp,i}$ 进行协整分析,得到独特的协整向量 $\mathbf{B}_{specific,i}$ 以及如式(13)所示的独特的协整特征:

[0058] $\gamma_{specific,i} = \mathbf{X}_{ns\perp,i} \mathbf{B}_{specific,i} \quad (13)$

[0059] 其中, $\mathbf{B}_{specific,i}$ 表示第 i 个故障类的独特的协整向量,向量个数由Johansen检验确

定。

[0060] (3.5) 获取独特的平稳特征矩阵: 求取每个故障类平稳的剩余独特特征矩阵和独特的协整特征矩阵的并集, 可以得到每个类独特的平稳特征矩阵:

$$[0061] \quad X_{\text{specific},i} = X_{s\perp,i} \cup \gamma_{\text{specific},i} \quad (14)$$

[0062] 其中, $X_{\text{specific},i}$ 包含了所有 (3.3) 中的平稳的剩余独特特征和 (3.4) 中的独特的协整特征, 统称为独特的平稳特征。

[0063] (3.6) 为每个故障类分别建立基于PCA的独特故障诊断模型, 该步骤包括以下子步骤:

[0064] (3.6.1) 选取上述得到的每个故障类的独特的平稳特征的样本作为总样本;

[0065] (3.6.2) 求取PCA的负载向量相当于求取总样本协方差矩阵的特征向量:

$$[0066] \quad \begin{aligned} S_{\Sigma,i} &= X_{\text{specific},i}^T X_{\text{specific},i} \\ S_{\Sigma,i} P_i &= \lambda P_i \end{aligned} \quad (15)$$

[0067] 其中, $S_{\Sigma,i}$ 表示总样本的协方差矩阵。这样得到的负载矩阵 P_i 就是PCA的最佳投影方向。

[0068] (3.6.3) 计算每类故障样本的主成分和故障衡量统计指标 $D_{\text{PCA},i,m}^2$:

$$[0069] \quad T_{\text{PCA},i} = X_{\text{specific},i} P_i$$

$$[0070] \quad D_{\text{PCA},i,m}^2 = (T_{\text{PCA},i} - \bar{t}_{\text{PCA},i})^T \Sigma_{\text{PCA},i}^{-1} (T_{\text{PCA},i} - \bar{t}_{\text{PCA},i}) \quad (16)$$

[0071] 其中, $T_{\text{PCA},i}$ 表示主成分矩阵, 包含了大部分的样本波动信息, $\bar{t}_{\text{PCA},i}$ 是 $T_{\text{PCA},i}$ 的均值向量, $\Sigma_{\text{PCA},i}$ 是 $T_{\text{PCA},i}$ 的协方差矩阵。

[0072] (3.6.4) 建立每类故障特征衡量指标的控制限: 由于假设过程数据服从多元广义高斯分布, 且样本数量足够大, 故认为变量衡量指标 $D_{\text{PCA},i,m}^2$ 服从加权 x^2 分布, 则可依据带权重的 x^2 分布建立统计量的控制限 $Ctrl_{D_{\text{PCA},i,m}^2}$ 。

[0073] (4) 在线故障诊断: 当故障发生后, 需要检测发生故障的变量, 确定故障的类型, 该步骤包括以下子步骤:

[0074] (4.1) 按照步骤 (1) 获取新数据 $\tilde{x}_f (1 \times J)$, 根据步骤 (2.2) 到 (2.3), 可以得到新的公共平稳特征矩阵 $\tilde{x}_{\text{new}} (1 \times J_{\text{alls}})$;

[0075] (4.2) 按照步骤 (2.4) 提取出关键故障特征, 得到新数据的关键故障特征样本 $\tilde{x}_{\text{key}} (1 \times J_{\text{key}})$;

[0076] (4.3) 调用公共故障诊断模型:

$$[0077] \quad \begin{aligned} \tilde{t}_{\text{new}} &= \tilde{x}_{\text{key}} R^* \\ \tilde{D}_{\text{new},i}^2 &= (\tilde{t}_{\text{new}} - \bar{t}_i)^T \Sigma_i^{-1} (\tilde{t}_{\text{new}} - \bar{t}_i) \end{aligned} \quad (17)$$

[0078] 其中, $R^* (J_{\text{key}} \times R_{\text{FDA}})$ 是步骤 (2.5.4) 中计算得到的FDA的广义特征向量矩阵, $\tilde{x}_{\text{key}}$ 是新数据的关键故障特征, $\tilde{t}_{\text{new}}$ 是 $\tilde{x}_{\text{key}}$ 的判别成分, $\bar{t}_i$ 和 Σ_i 均由步骤 (2.5.5) 求得, $\tilde{D}_{\text{new},i}^2$ 为公共故障衡量指标;

[0079] (4.4) 按步骤(3.2)将新的样本投影到剩余的独特子空间中,得到 $\tilde{\mathbf{x}}_{\perp, new}$ 。将 $\tilde{\mathbf{x}}_{\perp, new}$ 代入到每个故障类的独特协整模型中,可以得到新数据的独特平稳特征样本 $\tilde{\mathbf{x}}_{\perp, new, i}$;

[0080] (4.5) 调用独特故障诊断模型:

$$\tilde{\mathbf{t}}_{\perp, new, i} = \tilde{\mathbf{x}}_{\perp, new, i} \mathbf{P}_i$$

[0081]

$$\tilde{D}_{\perp, new, i}^2 = (\tilde{\mathbf{t}}_{\perp, new, i} - \bar{\mathbf{t}}_{PCA, i})^T \Sigma_{PCA, i}^{-1} (\tilde{\mathbf{t}}_{\perp, new, i} - \bar{\mathbf{t}}_{PCA, i}) \quad (18)$$

[0082] 其中, $\mathbf{P}_i$ 是步骤(3.6.2)中计算得到的PCA的负载矩阵, $\tilde{\mathbf{x}}_{\perp, new, i}$ 是新数据的独特平稳特征样本, $\tilde{\mathbf{t}}_{\perp, new, i}$ 是 $\tilde{\mathbf{x}}_{\perp, new, i}$ 的主成分矩阵, $\bar{\mathbf{t}}_{PCA, i}$ 和 $\Sigma_{PCA, i}$ 均由步骤(3.6.3)求得, $\tilde{D}_{\perp, new, i}^2$ 为独特故障指标;

[0083] (4.6) 对每类候选故障分别比较 $\tilde{D}_{new, i}^2$ 和其控制限 $Ctrl_{D_{l, m}}^2$, $\tilde{D}_{\perp, new, i}^2$ 和其控制限 $Ctrl_{D_{PCA, i, m}}^2$, 若新来的样本在该类候选故障的故障诊断模型下均没有超限报警, 则表明机组发生了该候选故障。

[0084] 与现有技术相比, 本发明的有益效果是: 本方法从百万千瓦超超临界机组的非平稳故障特性着手, 克服了传统多元统计分析方法无法用于非平稳故障过程的问题, 并充分挖掘了不同的故障类间的公共信息和每个故障类的独特信息, 使得方法可以应用于非平稳、高耦合、故障多样的复杂热力过程。该方法利用CA模型将非平稳故障过程转化为平稳故障过程, 然后综合分析了故障间的公共协整关系和每个故障内独特的故障特性, 建立了双重的故障诊断模型, 能更准确地区分生产过程中故障类别, 提高了在线故障诊断的性能, 有助于电厂维修人员准确地定位和修复故障, 确保了电厂生产的安全可靠运行并为大型燃煤机组的故障诊断方法的研究指明了新的方向。

附图说明

[0085] 图1是本发明基于互信息的多变量排序算法流程图;

[0086] 图2是本发明结合协整分析和双重故障模型的在线故障诊断方法的流程图;

[0087] 图3是本发明调用故障#1的双重诊断模型针对故障#1的在线故障诊断结果图, (a) 为公共故障诊断模型, (b) 为独特故障诊断模型;

[0088] 图4是本发明调用故障#2的双重诊断模型针对故障#1的在线故障诊断结果图, (a) 为公共故障诊断模型, (b) 为独特故障诊断模型;

[0089] 图5是本发明调用故障#8的双重诊断模型针对故障#1的在线故障诊断结果图, (a) 为公共故障诊断模型, (b) 为独特故障诊断模型;

[0090] 图6是分别调用故障#4的未进行非平稳处理的FDA故障诊断模型和双重诊断模型针对故障#4的在线故障诊断结果对比图, (a) 为未进行非平稳处理的模型, ; (b) 为公共故障诊断模型; (c) 为独特故障诊断模型;

[0091] 图7是分别调用故障#5的未进行非平稳处理的FDA故障诊断模型和双重诊断模型针对故障#4的在线故障诊断结果对比图, (a) 为未进行非平稳处理的模型, ; (b) 为公共故障诊断模型; (c) 为独特故障诊断模型;

[0092] 图8是分别调用故障#7的未进行非平稳处理的FDA故障诊断模型和双重诊断模型针对故障#4的在线故障诊断结果对比图, (a) 为未进行非平稳处理的模型,; (b) 为公共故障诊断模型; (c) 为独特故障诊断模型。

具体实施方式

[0093] 下面结合附图和具体实例,对本发明作进一步详细说明。

[0094] 大型燃煤发电的生产过程,尤其是故障过程,是典型的非平稳过程,其中一部分变量具有明显的非平稳特性,如凝汽器循环水压力、高压加热器进汽压力、除氧器进口凝结水量等。本发明以浙能集团下属嘉华电厂7号机组为例,该机组的功率为10000MW,为百万千瓦超超临界机组,包括46个过程变量,这些变量涉及到压力、温度、流量、流速等。

[0095] 如图2所示,本发明是一种面向百万千瓦超超临界机组非平稳故障特性的在线故障诊断方法,包括以下步骤:

[0096] (1) 获取过程分析数据:设一个百万千瓦超超临界机组热力系统具有J个过程变量,则每一次采样可以得到一个 $1 \times J$ 的向量,采样N次后得到一个二维矩阵 $X(N \times J)$ 。本实例中,采样周期为1分钟,每个故障类型采集1000个样本,46个过程变量,所测变量为运行过程中的流量、温度、速率、给料量、阀门开度等;分别获取8组不同类型的故障数据二维矩阵,统一表示为 $X_i(1000 \times J)$,其中下标i表示故障的类别;

[0097] (2) 建立公共的故障诊断模型,该步骤通过以下子步骤来实现:

[0098] (2.1) 平稳性检验:分别对每个故障类的所有变量进行ADF检验,得到的平稳变量矩阵表示为 $X_{s,i}(N \times J_{s,i})$,非平稳变量矩阵表示为 $X_{ns,i}(N \times J_{ns,i})$,其中,下标s表示平稳变量,ns表示非平稳变量,i表示故障的类别,括号里的 $J_{s,i}$ 表示平稳变量的个数, $J_{ns,i}$ 表示非平稳变量个数。

[0099] (2.2) 建立公共的协整模型,该步骤由以下子步骤来实现:

[0100] (2.2.1) 获取公共平稳变量矩阵:通过(2.1)区分了每个故障类中的平稳和非平稳变量,求取所有故障类中平稳变量的交集,可以得到对于每个故障类都平稳的变量,组成公共平稳变量矩阵:

$$[0101] \quad X_s(N \times J_s) = X_{s,1} \cap X_{s,2} \cap \dots \cap X_{s,C} \quad (1)$$

[0102] 其中, $\mathbf{X}_{s,j} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_j, \dots, \mathbf{x}_{J_{s,j}}]$ $j \in (1, 2, \dots, J_{s,i})$, $\mathbf{x}_j$ 代表第i个故障类中的平稳变量, $\cap$ 表示求取不同矩阵的交集, J_s 表示综合C类故障后的公共平稳变量个数。

[0103] (2.2.2) 获取所有非平稳变量:求取所有故障类中非平稳变量的并集,可以得到8种故障类型下的全部非平稳变量集合:

$$[0104] \quad X_{ns}(N \times J_{ns}) = X_{ns,1} \cup X_{ns,2} \cup \dots \cup X_{ns,C} \quad (2)$$

[0105] 其中, $\mathbf{X}_{ns,j} = [\mathbf{x}_{n,1}, \mathbf{x}_{n,2}, \dots, \mathbf{x}_{n,j}, \dots, \mathbf{x}_{n,J_{ns,j}}]$ $j \in (1, 2, \dots, J_{ns,i})$, $\mathbf{x}_{n,j}$ 表示第i个故障类中的非平稳变量, $\cup$ 表示求取不同矩阵的并集, J_{ns} 表示综合C类故障后的所有非平稳变量个数。

[0106] (2.2.3) 建立公共协整模型:对(3.2)中的所有非平稳变量矩阵 $X_{ns}(N \times J_{ns})$ 进行协整分析,得到公共协整向量B以及如式(3)所示的公共协整特征:

$$[0107] \quad \gamma = X_{ns} B \quad (3)$$

[0108] 其中,共得到13个残差序列,用Johansen检验残差的平稳性,得到11个平稳的残差

序列,即得到11个公共协整向量,表示为 $B(J_{ns} \times R_{CA})$ 。 $\gamma = X_{ns}B$ 是从所有类的非平稳向量中提取的公共协整特征矩阵,表征了公共的故障特性。

[0109] (2.3) 获取公共平稳特征矩阵:求取公共平稳变量矩阵和公共协整特征矩阵的并集,可以得到公共的平稳特征矩阵:

$$[0110] \quad X_{alls}(N \times J_{alls}) = X_s \cup \gamma \quad (4)$$

[0111] 其中, X_{alls} 包含了所有(2.2.1)中的33个公共平稳变量和(2.2.3)中的11个公共平稳协整特征,共44个特征,统称为公共平稳特征。下面为方便表示,用 J_{alls} 来表示公共平稳特征的个数。

[0112] (2.4) 提取关键故障特征,该步骤由以下子步骤来实现:

[0113] (2.4.1) 假设所有的故障样本服从多元广义高斯分布。定义一个虚拟的故障变量 C ,可能的取值的个数与故障数据二维矩阵的组数相同,即为8,在这里用 C 表示。服从如下的概率分布 $P(C=c) = P(c) = \frac{1}{C}$ 。根据互信息准则, x_{alls} 和故障变量 C 间的互信息可以按下式表示

$$[0114] \quad \begin{aligned} I(x_{all}; C) &= \sum_{c=1}^C \int_x P(c, x) \log \left(\frac{P(c, x)}{P(c)P(x)} \right) dx \\ &= \sum_{c=1}^C \int_x P(c) P(x|c) \log \left(\frac{P(c)P(x|c)}{P(c)P(x)} \right) dx \\ &= \sum_{c=1}^C P(c) \int_x P(x|c) \log(P(x|c)) dx - \sum_{c=1}^C P(c) \int_x P(x|c) \log(P(x)) dx \end{aligned} \quad (5)$$

[0115] 其中, x_{alls} 是公共平稳特征矩阵中的一个随机向量, x 表示其中的一个样本,样本服从散布矩阵为 Σ ,形状参数为 β 的多元广义高斯分布。 x_{alls} 在 $C=c$ 条件下的概率密度服从参数为 Σ_c 和 β_c 。

[0116] (2.4.2) 将所有特征按互信息大小递减排序后,需要找出那些在所有故障类中分类性能最好的一组关键特征,表示为 X_{key} 。该关键特征 $X_{key} = [X_{key,1}, X_{key,2}, \dots, X_{key,i}, \dots, X_{key,C}]$, $X_{key,i}$ 中具有 N_i 个样本。

[0117] (2.5) 建立基于FDA的公共故障诊断模型,该步骤包括以下子步骤:

[0118] (2.5.1) 选取上述得到的所有故障类的关键特征中的样本作为总样本

$\mathbf{X}_{key} \left(\sum_{i=1}^k N_i \times J_{key} \right)$,其中, $\mathbf{X}_{key} \left(\sum_{i=1}^k N_i \times J_{key} \right)$ 由 $X_{key,i}$ ($i=1, 2, \dots, k$) 组成, J_{key} 代表关键故障特征的个数;

[0119] (2.5.2) 数据准备:分别计算总样本均值向量 $\bar{\mathbf{x}}$ 、每类故障样本均值向量 $\bar{\mathbf{x}}_i$,总类内散布矩阵 S_w 和类间的散布矩阵 S_b ,计算公式如下:

$$S_i = \sum_{x_i \in X_{key,i}} (x_i - \bar{x}_i)(x_i - \bar{x}_i)^T$$

$$[0120] \quad S_w = \sum_{i=1}^C S_i \quad (6)$$

$$S_b = \sum_{x_i \in X_{key,i}} N_i (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})^T$$

[0121] 其中, S_i 是每个故障类的散布矩阵;

[0122] (2.5.3) 最大化费舍尔准则: 求取使费舍尔准则函数最大的权重向量 w , 即求使类间离散度和类内离散度的比值最大的权重向量 w :

$$[0123] \quad J(w) = \frac{w^T S_b w}{w^T S_w w} \quad (7)$$

[0124] 其中, 权重向量 w 就是费舍尔最优判别方向。

[0125] (2.5.4) 求使费舍尔判据最大的权重向量相当于一个求取广义特征向量的特征值问题:

$$[0126] \quad S_b w = \lambda S_w w \quad (8)$$

[0127] 其中, 特征值 λ 代表故障类之间的总体可分程度。当 S_w 为非奇异阵时, 上式可以转化成常规的特征值问题:

$$[0128] \quad S_w^{-1} S_b w = \lambda w \quad (9)$$

[0129] 求解上式可以得到 R_{FDA} 个非零的广义特征值以及广义特征向量矩阵 $R^* (J_{key} \times R_{FDA})$, $R_{FDA} = C - 1$ 。

[0130] (2.5.5) 计算每类故障样本的类内判别成分和故障衡量统计指标 $D_{i,m}^2$:

$$[0131] \quad T_i = X_{key,i} R^*$$

$$[0132] \quad D_{i,m}^2 = (T_i - \bar{t}_i)^T \Sigma_i^{-1} (T_i - \bar{t}_i) \quad (10)$$

[0133] 其中, T_i 表示判别成分矩阵, $\bar{t}_i$ 是 T_i 的均值向量, Σ_i 是 T_i 的协方差矩阵。

[0134] (2.5.6) 建立每类故障特征衡量指标的控制限: 由于假设过程数据服从多元广义高斯分布, 且样本数量足够大, 故认为变量衡量指标 $D_{i,m}^2$ 服从加权 χ^2 分布, 则可依据带权重的 χ^2 分布建立统计量的控制限 $Ctrl_{D_{i,m}^2}$ 。

[0135] (3) 建立独特的故障诊断模型, 该步骤通过以下子步骤来实现:

[0136] (3.1) 利用式 (11) 提取剩余的协整关系:

$$[0137] \quad B^\perp = [I - B(B^T B)^{-1} B^T] \quad (11)$$

[0138] 其中, B 是从 (2.2.3) 中提取的公共协整向量矩阵, I 表示单位矩阵。 $B^\perp$ 是 B 的正交空间, 表示剩余的协整方向。

[0139] (3.2) 将每个故障类中非平稳变量 $\tilde{X}_{ns,i} (N \times J_{ns})$ 投影到剩余的独特子空间中, 得到剩余的独特特征:

$$[0140] \quad X_{\perp,i} = \tilde{X}_{ns,i} B^\perp \quad (12)$$

[0141] 其中, $\tilde{\mathbf{X}}_{ns,i}$ 表示 $\mathbf{X}_{ns}$ 中的第 i 类故障样本矩阵。

[0142] (3.3) 平稳性检验:

[0143] 分别对 (3.2) 得到的每个故障类中所有非平稳变量间剩余的独特特征 (以下简称每个故障类的剩余独特特征) 进行 ADF 检验, 得到的平稳的剩余独特特征矩阵表示为 $\mathbf{X}_{s\perp,i}$, 非平稳的剩余独特特征矩阵表示为 $\mathbf{X}_{ns\perp,i}$ 。

[0144] (3.4) 建立每个故障类独特的协整模型:

[0145] 对 (3.3) 中的非平稳的剩余独特特征矩阵 $\mathbf{X}_{ns\perp,i}$ 进行协整分析, 得到独特的协整向量 $\mathbf{B}_{specific,i}$ 以及如式 (13) 所示的独特的协整特征:

$$[0146] \quad \gamma_{specific,i} = \mathbf{X}_{ns\perp,i} \mathbf{B}_{specific,i} \quad (13)$$

[0147] 其中, $\mathbf{B}_{specific,i}$ 表示第 i 个故障类的独特的协整向量, 向量个数由 Johansen 检验确定。

[0148] (3.5) 获取独特的平稳特征矩阵: 求取每个故障类平稳的剩余独特特征矩阵和独特的协整特征矩阵的并集, 可以得到每个类独特的平稳特征矩阵:

$$[0149] \quad \mathbf{X}_{specific,i} = \mathbf{X}_{s\perp,i} \cup \gamma_{specific,i} \quad (14)$$

[0150] 其中, $\mathbf{X}_{specific,i}$ 包含了所有 (3.3) 中的平稳的剩余独特特征和 (3.4) 中的独特的协整特征, 统称为独特的平稳特征。

[0151] (3.6) 为每个故障类分别建立基于 PCA 的独特故障诊断模型, 该步骤包括以下子步骤:

[0152] (3.6.1) 选取上述得到的每个故障类的独特的平稳特征的样本作为总样本;

(3.6.2) 求取 PCA 的负载向量相当于求取总样本协方差矩阵的特征向量:

$$[0153] \quad \begin{aligned} \mathbf{S}_{\Sigma,i} &= \mathbf{X}_{specific,i}^T \mathbf{X}_{specific,i} \\ \mathbf{S}_{\Sigma,i} \mathbf{P}_i &= \lambda \mathbf{P}_i \end{aligned} \quad (15)$$

[0154] 其中, $\mathbf{S}_{\Sigma,i}$ 表示总样本的协方差矩阵。这样得到的负载矩阵 $\mathbf{P}_i$ 就是 PCA 的最佳投影方向。

[0155] (3.6.3) 计算每类故障样本的主成分和故障衡量统计指标 $D_{PCA,i,m}^2$:

$$[0156] \quad \mathbf{T}_{PCA,i} = \mathbf{X}_{specific,i} \mathbf{P}_i$$

$$[0157] \quad D_{PCA,i,m}^2 = (\mathbf{T}_{PCA,i} - \bar{\mathbf{t}}_{PCA,i})^T \boldsymbol{\Sigma}_{PCA,i}^{-1} (\mathbf{T}_{PCA,i} - \bar{\mathbf{t}}_{PCA,i}) \quad (16)$$

[0158] 其中, $\mathbf{T}_{PCA,i}$ 表示主成分矩阵, 包含了大部分的样本波动信息, $\bar{\mathbf{t}}_{PCA,i}$ 是 $\mathbf{T}_{PCA,i}$ 的均值向量, $\boldsymbol{\Sigma}_{PCA,i}$ 是 $\mathbf{T}_{PCA,i}$ 的协方差矩阵。

[0159] (3.6.4) 建立每类故障特征衡量指标的控制限; 由于假设过程数据服从多元广义高斯分布, 且样本数量足够大, 故认为变量衡量指标 $D_{PCA,i,m}^2$ 服从加权 χ^2 分布, 则可依据带权重的 χ^2 分布建立统计量的控制限 $Ctr_{D_{PCA,i,m}^2}$ 。

[0160] (4) 在线故障诊断: 当故障发生后, 需要检测发生故障的变量, 确定故障的类型, 该步骤包括以下子步骤:

[0161] (4.1) 按照步骤 (1) 获取新数据 $\tilde{\mathbf{x}}_f (1 \times J)$, 根据步骤 (2.2) 到 (2.3), 可以得到新的公共平稳特征矩阵 $\tilde{\mathbf{x}}_{new} (1 \times J_{alls})$;

[0162] (4.2) 按照步骤 (2.4) 提取出关键故障特征, 得到新数据的关键故障特征样本 $\tilde{\mathbf{x}}_{key} (1 \times J_{key})$;

[0163] (4.3) 调用公共故障诊断模型:

$$\tilde{\mathbf{t}}_{new} = \tilde{\mathbf{x}}_{key} \mathbf{R}^*$$

[0164]

$$\tilde{D}_{new,i}^2 = (\tilde{\mathbf{t}}_{new} - \bar{\mathbf{t}}_i)^T \Sigma_i^{-1} (\tilde{\mathbf{t}}_{new} - \bar{\mathbf{t}}_i) \quad (17)$$

[0165] 其中, $\mathbf{R}^* (J_{key} \times R_{FDA})$ 是步骤 (2.5.4) 中计算得到的FDA的广义特征向量矩阵, $\tilde{\mathbf{x}}_{key}$ 是新数据的关键故障特征, $\tilde{\mathbf{t}}_{new}$ 是 $\tilde{\mathbf{x}}_{key}$ 的判别成分, $\bar{\mathbf{t}}_i$ 和 Σ_i 均由步骤 (2.5.5) 求得, $\tilde{D}_{new,i}^2$ 为公共故障衡量指标;

[0166] (4.4) 按步骤 (3.2) 将新的样本投影到剩余的独特征子空间中, 得到 $\tilde{\mathbf{x}}_{\perp, new}$ 。将 $\tilde{\mathbf{x}}_{\perp, new}$ 代入到每个故障类的独特协整模型中, 可以得到新数据的独特平稳特征样本 $\tilde{\mathbf{x}}_{\perp, new,i} (1 \times J_{\perp})$;

[0167] (4.5) 调用独特故障诊断模型:

$$\tilde{\mathbf{t}}_{\perp, new,i} = \tilde{\mathbf{x}}_{\perp, new,i} \mathbf{P}_i$$

$$\tilde{D}_{\perp, new,i}^2 = (\tilde{\mathbf{t}}_{\perp, new,i} - \bar{\mathbf{t}}_{PCA,i})^T \Sigma_{PCA,i}^{-1} (\tilde{\mathbf{t}}_{\perp, new,i} - \bar{\mathbf{t}}_{PCA,i}) \quad (18)$$

[0170] 其中, $\mathbf{P}_i$ 是步骤 (3.6.2) 中计算得到的PCA的负载矩阵, $\tilde{\mathbf{x}}_{\perp, new,i}$ 是新数据的独特平稳特征样本, $\tilde{\mathbf{t}}_{\perp, new,i}$ 是 $\tilde{\mathbf{x}}_{\perp, new,i}$ 的主成分矩阵, $\bar{\mathbf{t}}_{PCA,i}$ 和 $\Sigma_{PCA,i}$ 均由步骤 (3.6.3) 求得, $\tilde{D}_{\perp, new,i}^2$ 为独特故障指标;

[0171] (4.6) 对每类候选故障分别比较 $\tilde{D}_{new,i}^2$ 和其控制限 $Ctrl_{D_{i,m}}^2$, $\tilde{D}_{\perp, new,i}^2$ 和其控制限 $Ctrl_{D_{PCA,i,m}}^2$, 若新来的样本在该类候选故障的故障诊断模型下均没有超限报警, 则表明机组发生了该候选故障。

[0172] 本发明通过分析百万千瓦超超临界机组故障过程的非平稳特性和故障间的关系, 根据历史故障数据建立双重故障诊断模型, 工程师可以实时获得新过程采样数据的在线故障诊断结果, 判断故障类别从而进行相应的故障修复措施, 消除故障影响。当基于历史故障数据建立的双重诊断模型都能很好地容纳新故障样本所求得的指标, 即调用该类故障的新样本指标既在公共故障诊断模型建立的控制限内, 也在独特故障诊断模型建立的控制限内, 则说明新样本属于该类故障, 那么当前新样本所属故障就被正确诊断出来。由于一个故障样本可能被多个候选故障类型的故障诊断模型所解释, 我们在此计算两个指标来判断本发明的故障诊断性能, 分别是正确分类率 (CCR) 和错误分类率 (FCR)。

$$\begin{aligned} \text{正确分类率} &= \frac{\text{故障}i\text{控制限以内的样本数}}{\text{没有超限的样本总数}} \\ \text{错误分类率} &= \frac{\text{故障}i\text{控制限以外的样本数}}{\text{实际样本总数}} \end{aligned} \quad (19)$$

[0174] 故障 i 的正确分类率越大, 说明故障 i 的故障诊断模型能更好地解释这类故障样本, 即该类故障样本属于故障 i 的可能性越大。计算错误分类率则可以直观地判断有这类故

障样本中有多少样本不能被正确地识别。针对本发明的双重模型,某个样本只有同时在候选故障的公共模型和独特模型的控制限内,才被计入没有超限的样本数。表1总结了本发明所提出的方法与未进行非平稳处理,直接使用基于变量选择的FDA方法针对8种故障类型的故障诊断性能指标的比较结果。

[0175] 表1三种方法故障诊断性能对比(衡量指标:正确分类率和错误分类率)

故障#	方法					
	未进行非平稳处理的 FDA 方法		基于公共故障诊断模型的方法		基于双重故障诊断模型的方法	
	CCR	FCR	CCR	FCR	CCR	FCR
1	47.93	3.33	70.92	0	99.00	0
2	60.18	0	87.05	0	100	0
3	57.25	2.67	89.02	0	100	0
4	75.56	1.33	100	0.3	100	0.3
5	14.32	23.33	14.66	16.67	54.15	2.33
6	14.48	22.67	15.02	18	58.78	5
7	25	2	46.52	0	75.65	0
8	86.42	0	100	0	100	0

[0177] 从表1中可以看出,基于本发明方法的故障诊断方法提高了实际在线故障诊断的可靠性和可信度,对于8种故障的诊断均有较高的正确率,其性能优于传统方法。此外,图3(a)(b)至图5(a)(b)展示了本发明所提出的方法对于故障1的在线诊断结果,虚线表示每类故障基于指标的控制限,点线表示指标。可以看出,故障1的样本只有在调用故障#1的诊断模型时,同时处于公共故障诊断模型和独特故障诊断模型控制限以下,说明本发明能正确地诊断出故障1,说明了本发明方法的有效性。图6(a)(b)(c)至图8(a)(b)(c)对比了本发明方法和未进行非平稳处理直接使用FDA方法的性能。从图6(a)、7(a)、8(a)可以看出,属于故障#4的样本可能同时在其他故障的控制限以下,导致较大的故障误判率。与之对比的图6(b)(c)、图7(b)(c)、图8(b)(c)可以看出,属于故障#4的样本只在调用故障#4建立的双重诊断模型的控制限以下,即该发明能正确诊断出故障#4,说明了本发明方法的有效性。总体说来,本发明方法在针对百万千瓦超超临界机组这一典型非平稳多故障过程的在线故障诊断具有优越性,可以帮助现场工程师准确有效地诊断并及时修复故障,保证了实际生产过程的安全性和可靠性。

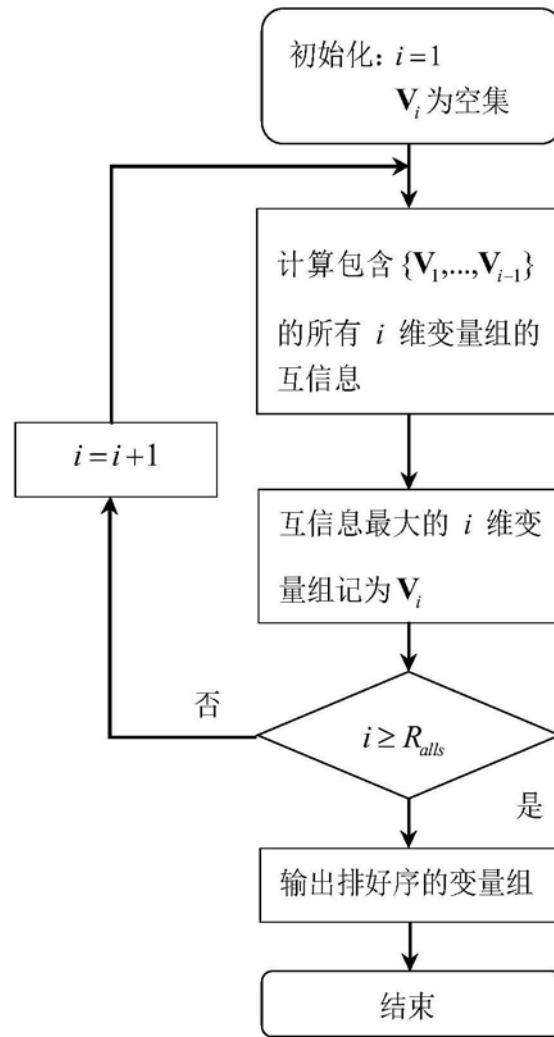


图1

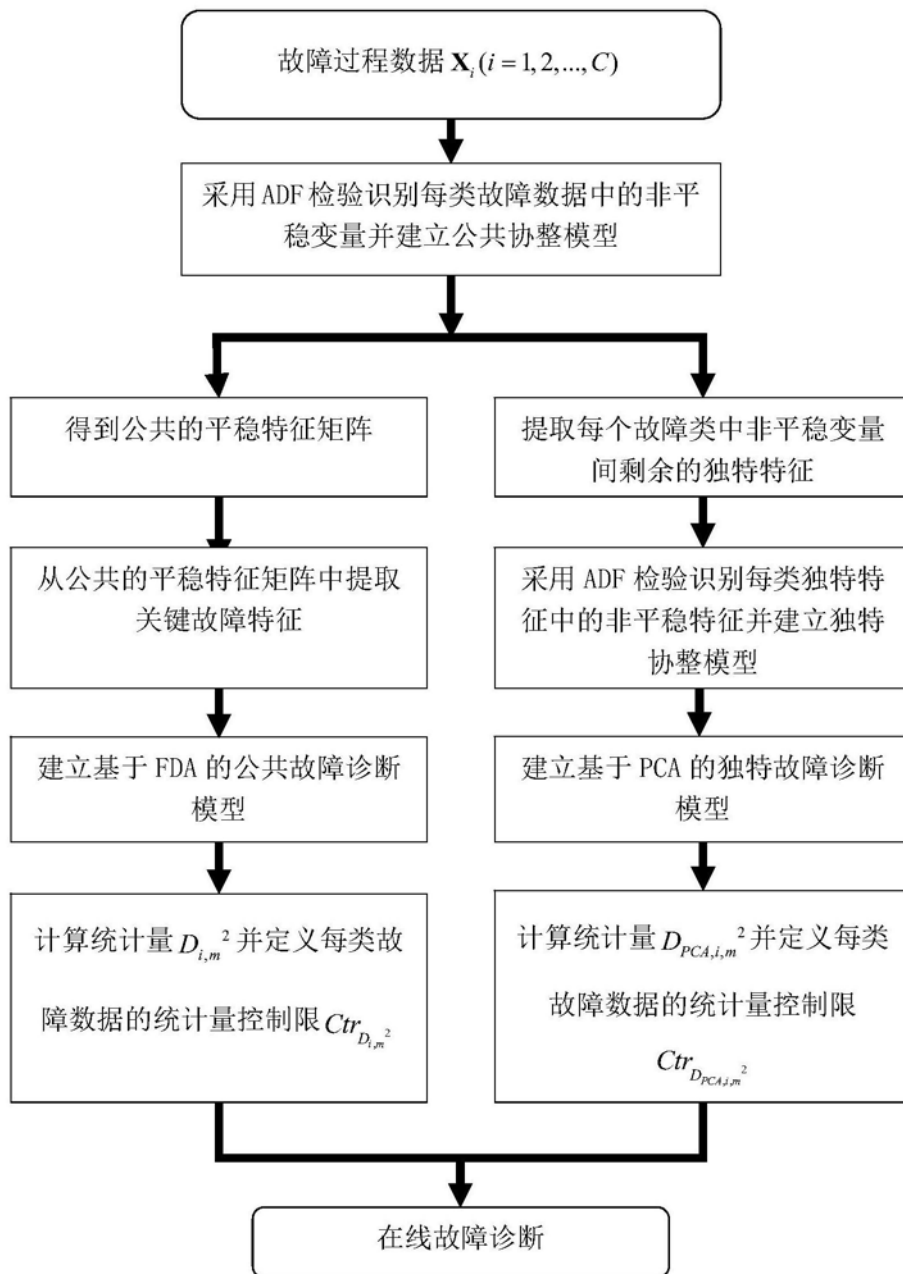


图2

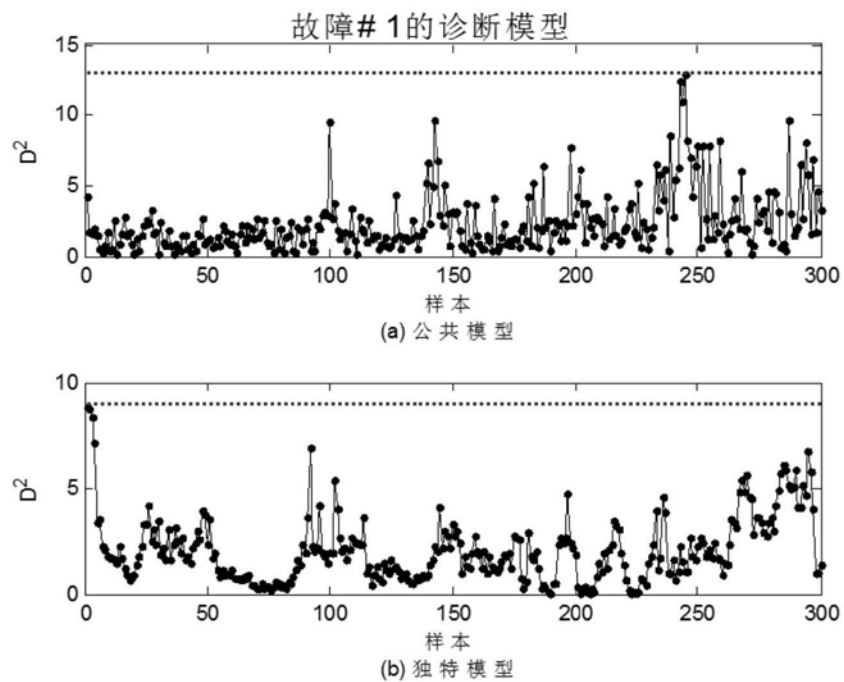


图3

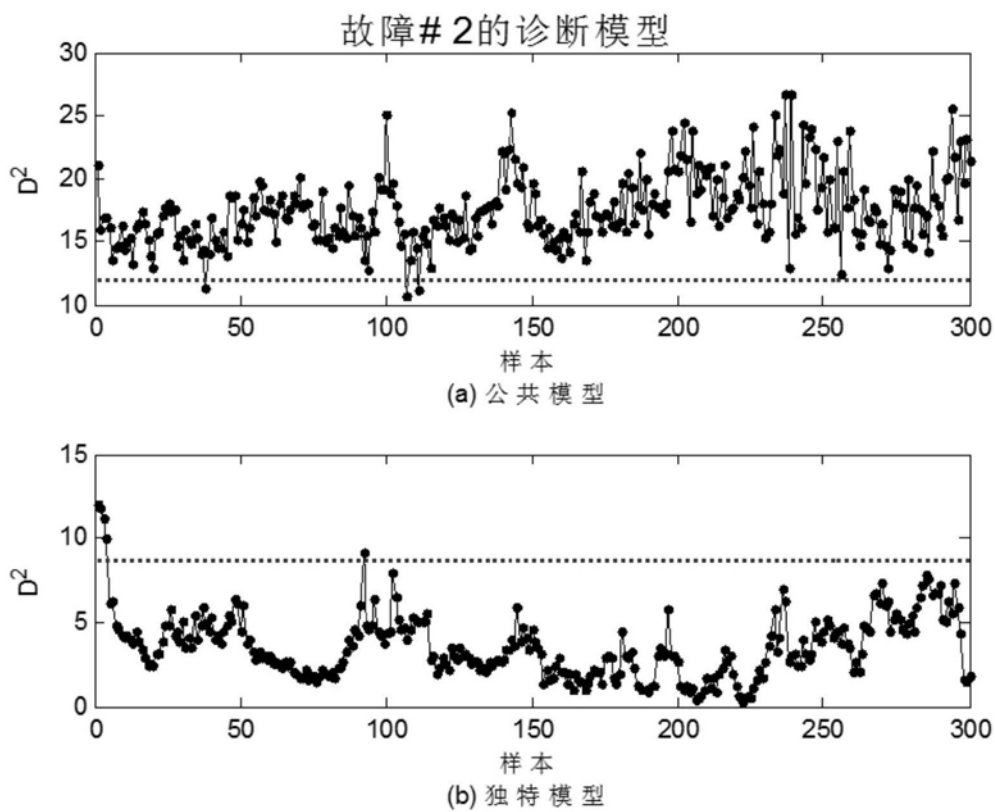


图4

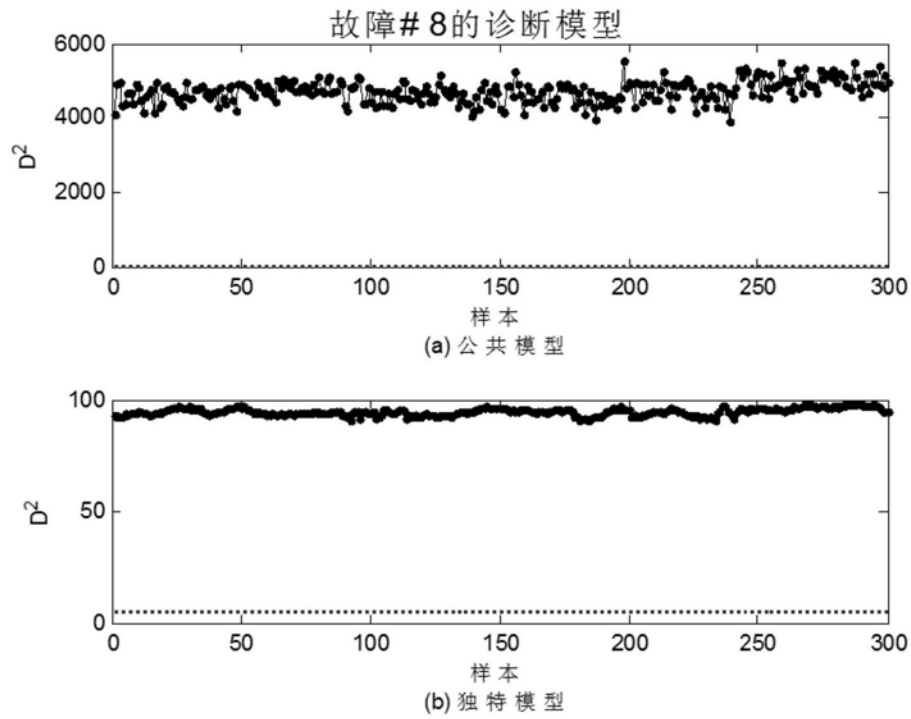


图5

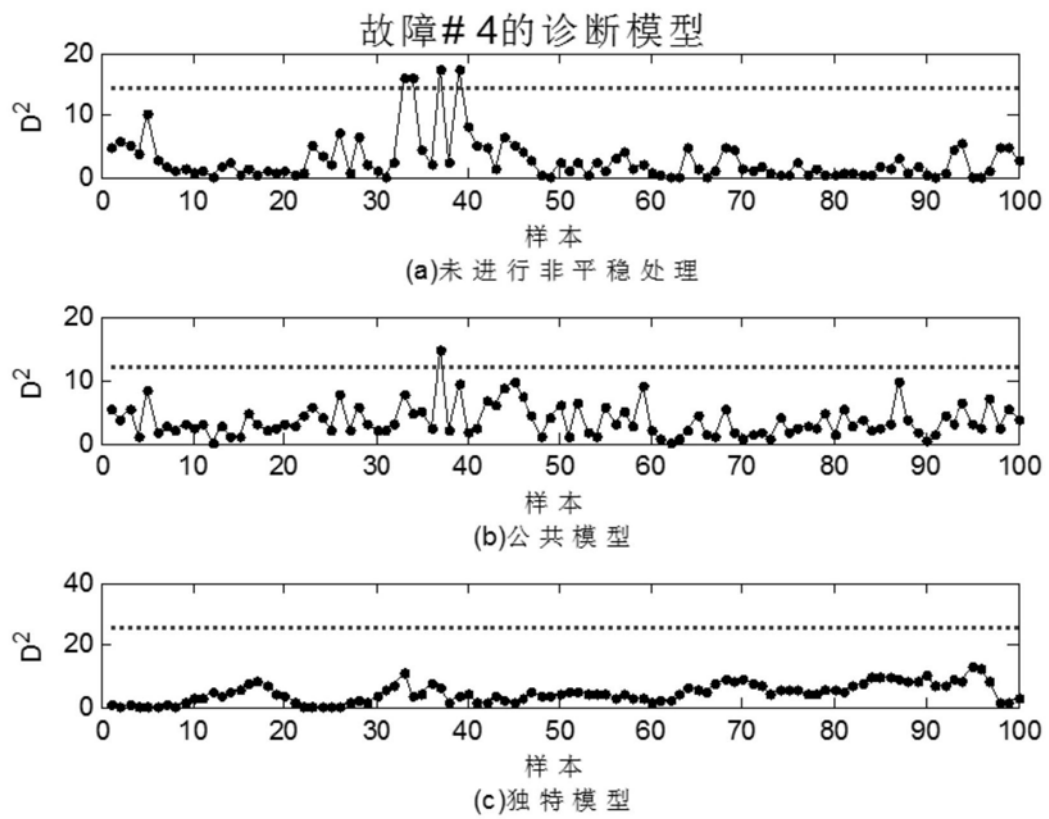


图6

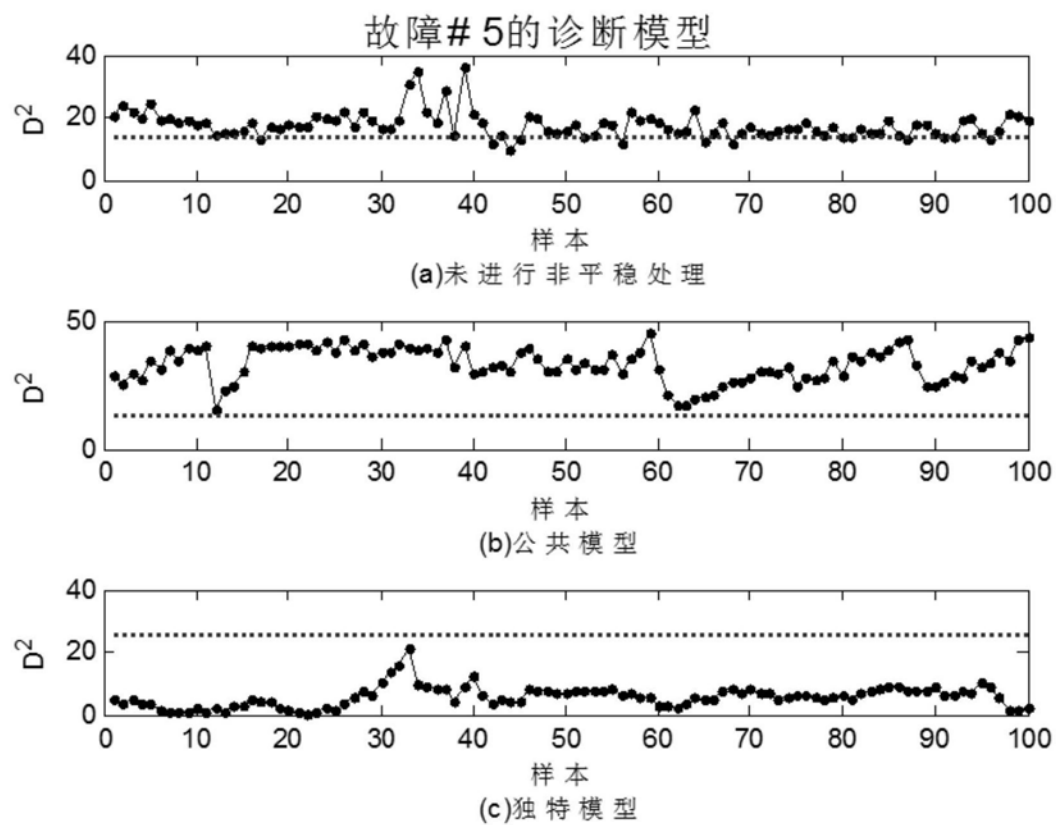


图7

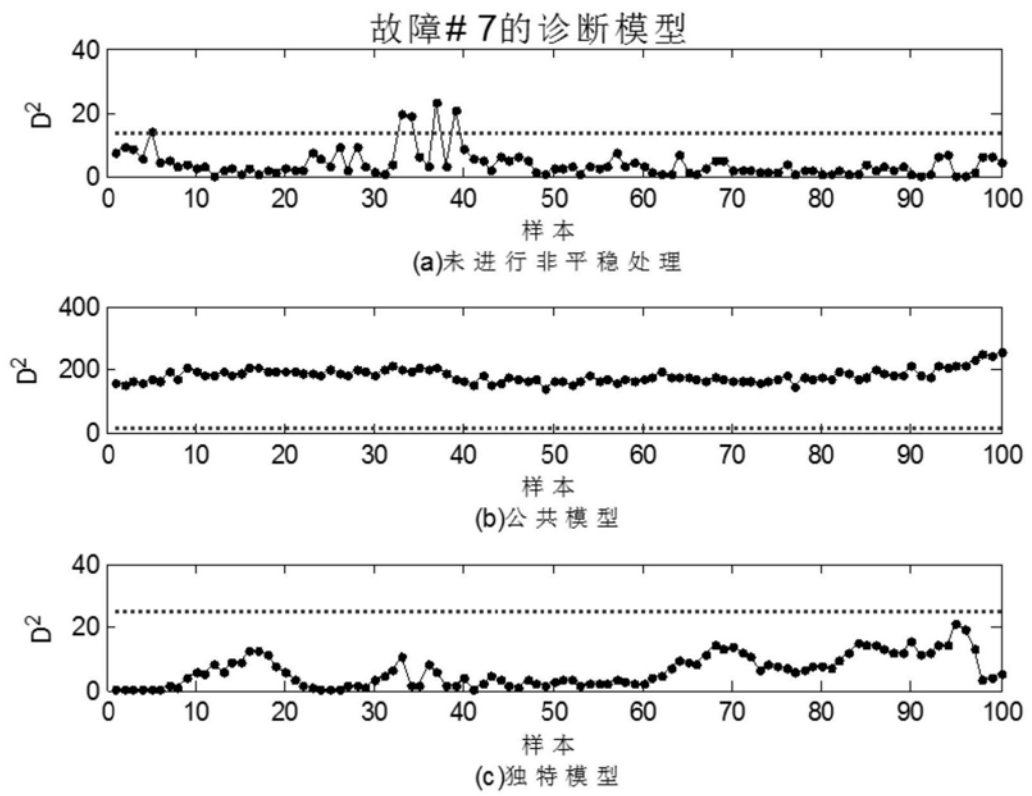


图8